PASSPOINT SETTING FOR MULTIPLE CHOICE EXAMINATIONS

INTRODUCTION

An examination is a tool for measuring individual characteristics that will allow the exam administrator to make decisions regarding the individual test takers based on their possession of those characteristics. An agency using a multiple choice examination to make employment decisions or determinations of which candidates should receive a license or certification must determine a threshold of acceptable performance on the score continuum, in other words, who should pass the test. This threshold is often referred to as a passpoint, cut score or cutoff score. Because of the implications and consequences of the decisions made based on exam results, agencies have a responsibility to use great care and consideration when setting a passpoint.

The primary purpose of this paper is to outline the considerations that agencies should take into account when setting a passpoint for a multiple choice examination. Some factors that need to be considered are:

- Whether the examination will be used as a pass/fail device or will be used to rank candidates.
- The purpose of the examination (employment, certification/licensure, etc.).
- Whether the criteria used to set the passpoint is norm-referenced or criterion-referenced.
- When the passpoint should be established (before or after the test administration).
- If the passpoint adversely impacts any protected groups.
- Administrative issues (such as the number of candidates, the number of job openings, the length of time a subsequent eligibility list will remain active, etc.).

According to the *Principles for the Validation and Use of Personnel Selection Procedures* from the Society for Industrial and Organizational Psychology (SIOP):

> *"There is no single method for establishing cutoff scores. If based on valid predictors demonstrating linearity or monotonicity throughout the range of prediction, cutoff scores may be set as high or as low as needed to meet the requirements of the organization (SIOP, 2003, p. 46)."*

Although agencies may be able to set examination passpoints where they best suit their needs, the passpoint-setting process is not without its pitfalls. Most legal challenges of testing programs involve allegations of incorrect or unfair decisions made as a result of applying a particular passpoint *(Cizek, 2006)*.

An agency using a multiple choice examination to make employment decisions or determinations of which candidates should receive a license or certification must determine a threshold of acceptable performance on the score continuum, in other words, who should pass the test.

CPS HR CONSULTING

Any agency that uses multiple choice examinations to assess applicants for jobs, licenses or credentials needs to determine a passpoint for the examination that:

(1) is legally defensible in terms of validity,

(2) minimizes adverse impact against protected groups and

(3) allows a sufficient number of candidates to pass or not pass in order to obtain the desired number of candidates to either be placed on the eligibility list or to proceed with additional phases of the assessment process. Some of the more significant considerations for testing agencies in regard to setting passpoints for multiple choice examinations are discussed below.

ISSUES ASSOCIATED WITH SETTING MULTIPLE CHOICE EXAMINATION PASSPOINTS

False Positives vs. False Negatives. There are two types of errors associated with setting examination passpoints: false positive results and false negative results. False positives are those candidates that pass an examination without possessing sufficient expertise to safely and successfully perform the duties associated with a job, or in a field of practice for which the candidate is seeking licensure or certification. False negatives are those candidates that do possess the requisite expertise, yet do not pass the examination. Agencies must consider the intended use of the test scores and be aware of the impact of potential false positive and false negative results in their passpoint-setting process.

Administrative Considerations. For most agencies, decisions on where to set a passpoint are heavily influenced by administrative considerations, such as the number of anticipated job vacancies, the number of candidates needed on an eligibility list, the duration of the eligible list and the costs associated with the assessment process. For example, a process resulting in too few candidates on an eligibility list must be continually re-administered, costing the organization additional resources. Administrative and economic issues are valid reasons that can be used to make passpoint decisions. However, these issues should be considered along with other factors such as potential adverse impact and the overall validity of the testing process.

Adverse Impact. It is not uncommon to set a passpoint that eliminates or minimizes the amount adverse impact on protected groups in accordance with civil rights law such as Title VII of the Civil Rights Act of 1964, the Age Discrimination in Employment Act of 1967, and the Civil Rights Act of 1991. Typically, this is accomplished by setting the passpoint so that the Four-Fifths Rule[1] is not violated for any protected groups. However, agencies must be cautious that they do not set the passpoint so low as to allow unqualified candidates to be placed on the eligible list. This concern is expressed in the *Uniform Guidelines on Employee Selection Procedures* (1978):

> "***Cutoff Scores.*** Where cutoff scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force. Where applicants are ranked on the basis of properly validated selection procedures and those applicants scoring below a higher cutoff score than appropriate in light of such expectations have little or no chance of being selected for employment, the higher cutoff score may be appropriate, but the degree of adverse impact should be considered."

Use of Natural Breaks in the Score Distribution. Some agencies set passpoints at natural breaks in the score distribution. The logic behind this practice is that it (1) provides a natural break between the higher and lower scoring candidates, indicating a larger difference in knowledge between the groups of candidates, and (2) it minimizes the ability of the lower scoring candidates to successfully protest their examination results. In theory, the larger the break between a candidate's non-passing score and the score needed to pass an examination, the less likely the candidate will be to protest the results of individual test questions (to try to gain enough extra points to achieve a score at or above the pass point). However, agencies should be aware that these natural breaks may not coincide with performance standards, and being based on the qualities of a particular candidate group, the passpoint can fluctuate from administration to administration.

Criterion-Referenced Passpoints. Criterion-referenced passpoints are based on pre-established levels of competency, and are typically set prior to test administration. Criterion-referenced passpoints are frequently used in employment, licensure and certification testing. The primary purpose of licensure examinations is public protection (i.e., to ensure that the recipients of the license possess the minimum competency needed to successfully perform the job or service). Hence, with licensure examinations the preset passpoint is typically set at a level that demonstrates minimal competency on the part of the candidate.

The primary purpose of certification examinations, however, is to ensure that the candidate has achieved a degree of mastery of the subject area. The mastery level is above the minimally competent level, hence, certification examinations tend to have a higher passpoint than licensure examinations on the same subject matter. Standard 11.16 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) reads:

> *"The level of performance required for passing a credentialing test should depend on the knowledge and skills necessary for credential worthy performance in the occupation or profession and should not be adjusted to control the number or proportion of persons passing the test."*

The primary purpose of an employment examination is to predict job performance. Passpoints for employment examinations are therefore typically set at the point on the score continuum that represents minimal expected job performance but then may be adjusted based issues discussed above such as adverse impact and administrative considerations.

The most common method used to establish performance criteria for employment, licensure and certification examinations was developed by William Angoff (1971). When using the Angoff method to set a passpoint, Subject Matter Experts (SMEs) are asked to estimate the percentage of minimally-competent candidates that would answer each test question correctly. The SMEs' estimates for each test questions are averaged, producing an Angoff score for each item. When the Angoff scores for all the items in a test are averaged, the result is an estimate of the score that a minimally-competent candidate would obtain on the examination. Hence, the passpoint is typically set at or near this level of minimal competency.

Another method used to establish performance criteria is the Nedelsky Method *(Nedelsky, 1954).* When using the Nedelsky Method, SMEs are asked to determine the likelihood that borderline (i.e., minimally competent) candidates would be able to rule out each incorrect multiple choice alternative. A formula is then applied to assign a numerical score to each test question. The sum of these numerical scores is used to establish the passpoint for the examination.

CPS HR CONSULTING

Norm-Referenced Passpoints. Norm-referenced passpoints are based on the characteristics of the score distribution, such as range of scores, size of the standard deviation and placement of certain groups, such as women and minorities, within the distribution. They are typically used in employment examinations, often when the test user desires a specific number of candidates to pass the examination. Norm-referenced passpoint criteria are typically set after the examination is administered. Basing the cutoff score on norm-referenced criteria is perfectly legal and often encouraged to ensure adverse impact is minimized. However, similar to using natural breaks in the scores, a norm-referenced passpoint is based solely on the exam performance of a particular candidate group, which may or may not possess the expected level of competence.

Pass/Fail vs. Ranking Candidates. Pass/Fail examinations are often criterion-referenced, which means that candidates are being assessed as to whether they can demonstrate a pre-established amount of knowledge or ability. Licensure and certification examination are assessed on a pass/fail basis. Licensure tests typically use a passpoint that ensures a minimal amount of competency, whereby passpoints for certification exams are typically higher as their goal is to ensure mastery of a subject area. Often times, employment examinations might be scored on a pass/fail basis when they are used as a screening device (i.e., hurdle) in an assessment process, which involves multiple assessment instruments and/or when the agency is trying to minimize any adverse impact associated with ranking the candidates.

Examinations that are used to select candidates for employment will typically rank candidates on an eligible list. For employment examinations, the employer must set passpoints that will allow a sufficient number of qualified applicants to be placed on the eligible list to fill all of the agency's vacancies over a reasonable length of time. However, agencies that use a top-down approach to hiring (i.e., those scoring highest on the exam moving to the next phase of the process) from a ranked eligible list should determine if this approach adversely impacts protected groups. According to the *Uniform Guidelines on Employee Selection Procedures* (1978),

> *"If the use of a validated procedure for ranking results in greater adverse impact than its use as a screening device, the evidence of validity and utility must be sufficient to warrant use of the procedures as a ranking device."*

70% Cutoff Written into Agency Rules/Ordinances. Many public agencies are required by rule or ordinance to employ a 70% passpoint (or something similar) for all employment examinations. Many of these rules and ordinances were set several years ago to ensure that candidates possess a minimum level of competency and as a means to curb the practice of nepotism that was rampant in many public agencies. A preset passpoint does not work well with a norm-referenced examination as the testing agency has no control over how many (or how few) candidates will actually pass the test and be placed on the eligibility list. With criterion-referenced examinations, the test developer may feel the need to select items for the examination that have a known difficulty level so as to build an examination that has an overall difficulty level that is aligned with the pre-established passpoint. This process requires the test developer to have pre-tested or previously administered the items for an indication of how they will perform.

CPS HR CONSULTING

Rescaling. The process of rescaling involves changing the range of a distribution of scores without changing the shape of the distribution or the relative distance between the individual scores in the distribution. Rescaling is often used in agencies that are required to have a specific passpoint (e.g., 70%) and where the distribution of the raw scores on an assessment instrument does not result in the desired number of candidates passing the examination. The easiest way to rescale an examination is to add or subtract a constant from the candidates' raw scores. However, the practice of subtracting points from the candidates' scores could lead to challenges from candidates and eventual litigation. The more common way to rescale candidates' scores is to first convert the scores to *Z scores* (with a mean of 0 and a standard deviation of 1). The *Z scores* are then converted to *T scores* by multiplying the Z score by a constant to get a desired standard deviation for the new distribution and then adding another constant to get a desired mean for the new distribution.

Standard Error of Measurement (SEM). Another common strategy when using a criterion-related approach is to set the passpoint within ± 1 SEM of the minimally acceptable criterion-related standard. An assumption of testing is that there is some amount of error inherent to the test. The SEM is an indication of the candidate variability on the examination attributable to random errors of measurement. This process is considered to be a valid method for controlling error in the testing process. Setting the passpoint one SEM above the minimally competent level allows the agency to better control for False Positive errors; setting the passpoint one SEM below the minimally competent level allows the agency to better control for False Negative errors. The importance of the SEM in setting passpoints is exemplified in the *Standards for Educational and Psychological Testing*; Standard 2.14 reads "Where passpoints are specified for selection or classification, the standard error of measurement should be reported in the vicinity of each passpoint (AERA, APA, & NCME, 1999).

CONCLUSION

Setting a valid, fair and appropriate passpoint is often the most legally contentious activity associated with test development and test administration processes. There are no hard and fast rules, laws or regulations regarding where to set passpoints for multiple choice examinations. Agencies should consider all of the issues listed above that are applicable to their agency and exam use and balance the legal, ethical, economic and administrative concerns when making their passpoint decisions.

CPS HR consultants have extensive experience in setting multiple choice examination passpoints that are valid and legally defensible, and we work with our clients to devise passpoint strategies that meet the specific needs of the individual agency.

CPS HR CONSULTING

REFERENCES

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, D.C.: Author.

Angoff, William H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndikes (Ed.), *Educational Measurement* (pp 508 – 600), Washington D.C.: American Council on Education.

Cizek, Gregory J. (2006). Standard Setting. In Downing, Steven M. & Haladyna, Thomas M. (Ed.), *Handbook of Test Development* (pp. 225 – 258). Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Equal Employment Opportunity Commission, Civil Service Commission, U.S. Department of Labor, & U.S. Department of Justice. (1978). *Uniform Guidelines on Employee Selection Procedures*, 43 C.F.R. § 38295.

Nedelsky, L. (1954). Absolute grading standards for objective tests. *Educational and Psychological Measurement, 14*, 3 – 19.

Society for Industrial and Organizational Psychology, (2003). *Principles for the Validation and Use of Personnel Selection Procedures*, 2003, SIOP: Bowling Green, OH

CPS HR CONSULTING